

## CLAIMS

1. A method for targeted-disruption of an arbitrary gene in the genome of a living organism, comprising the steps  
5 of:
  - A) providing information of the entire sequence of the genome of the living organism;
  - B) selecting at least one arbitrary region of the sequence;
  - 10 C) providing a vector comprising a sequence complementary to the selected region and a marker gene;
  - D) transforming the living organism with the vector; and
  - E) placing the living organism in a condition allowing  
15 homologous recombination.
2. The method according to Claim 1 wherein in step B), the region comprises at least two regions.
- 20 3. The method according to Claim 1, wherein the vector further comprises a promoter.
4. The method according to Claim 1 further comprising the step of detecting an expression product of the marker gene.  
25
5. The method according to Claim 5 wherein the marker gene is located in the selected region.
6. The method according to Claim 1, wherein the marker  
30 is located outside of the selected region.
7. The method according to Claim 1, wherein the genome is the genome of *Thermococcus kodakaraensis* KOD1.

8. The method according to Claim 1, wherein the genome has a sequence set forth in SEQ ID NO: 1 or 1087.
- 5 9. The method according to Claim 1, wherein the region comprises a sequence encoding at least one sequence selected from the group consisting of SEQ ID NOs: 2-341, 343-722, 724-1086, 1088-1468, 1470-1837 and 1839-2157.
- 10 10. A nucleic acid molecule having a sequence set forth in SEQ ID NO: 1 or 1087.
11. A nucleic acid molecule comprising at least eight contiguous nucleic acids of a sequence set forth in SEQ ID  
15 NO: 1 or 1087.
12. A nucleic acid molecule comprising a sequence encoding an amino acid sequence encoding at least one sequence selected from the group consisting of SEQ ID NOs: 2-341,  
20 343-722, 724-1086, 1088-1468, 1470-1837 and 1839-2157; or a sequence having 70 % homology thereto.
13. A nucleic acid molecule wherein when the reading frame of Table 2 is f-1, f-2 or f-3, the nucleic acid molecule  
25 has a sequence from the position of nucleic acid number (sense strand, start) of SEQ ID NO: 1 of Table 2, to the position of nucleic acid number (sense strand, stop) or a sequence having at least 70 % homology thereto, or when the reading frame of Table 2 is r-1, r-2 or r-3, the nucleic  
30 acid molecule has a sequence from the position of nucleic acid number (antisense strand, start) of SEQ ID NO: 1087 of Table 2, to the position of nucleic acid number (antisense strand, stop) or a sequence having at least 70 % homology

thereto.

14. A polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NO:  
5 2-341, 343-722, 724-1086, 1088-1468, 1470-1837 and 1839-2157, or a sequence having at least 70 % homology thereto.

15. A polypeptide comprising at least three contiguous  
10 amino acids of an amino acid sequence selected from the group consisting of SEQ ID NO: 2-341, 343-722, 724-1086, 1088-1468, 1470-1837 and 1839-2157, or a sequence having at least 70 % homology thereto.

15 16. A polypeptide comprising at least eight contiguous amino acids of an amino acid sequence selected from the group consisting of SEQ ID NO: 2-341, 343-722, 724-1086, 1088-1468, 1470-1837 and 1839-2157, or a sequence having at least 70 % homology thereto.

20 17. A polypeptide comprising at least three contiguous amino acids of an amino acid sequence selected from the group consisting of SEQ ID NO: 2-341, 343-722, 724-1086, 1088-1468, 1470-1837 and 1839-2157, or a sequence having at least 70 %  
25 homology thereto, wherein the polypeptide has biological activity.

18. The polypeptide according to Claim 17, wherein the biological activity comprises a function set forth in Table  
30 2.

19. A method of screening for a heat resistant protein, comprising the steps of:

- A) providing the entire sequence of the genome of a thermoresistant living organism;
- B) selecting at least one arbitrary region of the sequence;
- 5 C) providing a vector comprising a sequence complementary to the selected region and a gene encoding a candidate for the heat resistance protein;
- D) transforming the living organism with the vector;
- E) placing the thermoresistant living organism in a
- 10 condition allowing to cause homologous recombination;
- F) selecting the thermoresistant living organism in which homologous recombination has occurred; and
- G) assaying to identify the thermoresistant protein.
- 15 20. A kit for screening a thermoresistant protein, comprising:
- A) a thermoresistant living organism; and
- B) a vector comprising a sequence complementary to the selected region and a gene encoding a candidate for the
- 20 thermoresistant protein.
21. The kit according to Claim 20, further comprising an assay system for identifying the thermoresistant protein.
- 25 22. The kit according to Claim 20, wherein the thermoresistant living organism is hyperthermophilic bacteria.
23. The kit according to Claim 20, wherein the
- 30 thermoresistant living organism is *Thermococcus kodakaraensis* KOD1.
24. A biomolecule chip having at least one nucleic acid

molecule having at least eight contiguous or non-contiguous nucleotides of the sequence set forth in SEQ ID NOs: 1 or 1087, or a variant thereof located therein.

5 25. The biomolecule chip according to Claim 24, wherein the nucleic acid molecule or the variant thereof is located to cover the sequence set forth in SEQ ID NO: 1 or 1087.

26. The biomolecule chip according to Claim 24, wherein  
10 the nucleic acid molecule or the variant thereof comprises any open reading frame of the sequence set forth in SEQ ID NO: 1 or 1087.

27. The biomolecule chip according to Claim 24, wherein  
15 the nucleic acid molecule or the variant thereof comprises substantially all open reading frames of the sequence set forth in SEQ ID NO: 1 or 1087.

28. The biomolecule chip according to Claim 24, wherein  
20 the nucleic acid molecule or the variant thereof comprises a sequence encoding at least one sequence selected from the group consisting of SEQ ID NOs: 2-341, 343-722, 724-1086, 1088-1468, 1470-1837 and 1839-2157.

25 29. The biomolecule chip according to Claim 24, wherein the nucleic acid molecule or the variant thereof comprises substantially all sequences encoding sequences selected from the group consisting of SEQ ID NOs: 2-341, 343-722, 724-1086, 1088-1468, 1470-1837 and 1839-2157.

30 30. The biomolecule chip according to Claim 24, wherein the nucleic acid molecule or the variant thereof comprises at least eight contiguous nucleotide lengths of

substantially all sequences encoding sequences selected from the group consisting of SEQ ID NOs: 2-341, 343-722, 724-1086, 1088-1468, 1470-1837 and 1839-2157.

5    31.    The biomolecule chip according to Claim 24, wherein the nucleic acid molecule or the variant thereof comprises at least fifteen contiguous nucleotide lengths of substantially all sequences encoding sequences selected from the group consisting of SEQ ID NOs: 2-341, 343-722,  
10    724-1086, 1088-1468, 1470-1837 and 1839-2157.

32.    The biomolecule chip according to Claim 24, wherein the nucleic acid molecule or the variant thereof comprises at least thirty contiguous nucleotide lengths of  
15    substantially all sequences encoding sequences selected from the group consisting of SEQ ID NOs: 2-341, 343-722, 724-1086, 1088-1468, 1470-1837 and 1839-2157.

33.    The biomolecule chip according to Claim 24, wherein  
20    the nucleic acid molecule or the variant thereof, comprises substantially all sequences encoding sequences selected from the group consisting of SEQ ID NOs: 2-341, 343-722, 724-1086, 1088-1468, 1470-1837 and 1839-2157, or sequences with one or more amino acid substitution, addition and/or  
25    deletion thereto.

34.    The biomolecule chip according to Claim 24, wherein the nucleic acid molecule or the variant thereof, comprises at least eight contiguous nucleotide lengths of  
30    substantially all sequences encoding sequences selected from the group consisting of SEQ ID NOs: 2-341, 343-722, 724-1086, 1088-1468, 1470-1837 and 1839-2157, or sequences with one or more amino acid substitution, addition and/or

deletion thereto.

35. The biomolecule chip according to Claim 24, wherein when the reading frame of Table 2 is f-1, f-2 or f-3, the nucleic acid molecule or the variant thereof, has a sequence from the position of nucleic acid number (sense strand, start) of SEQ ID NO: 1 of Table 2, to the position of nucleic acid number (sense strand, stop) or a sequence having at least 70 % homology thereto, or when the reading frame of Table 2 is r-1, r-2 or r-3, the nucleic acid molecule has a sequence from the position of nucleic acid number (antisense strand, start) of SEQ ID NO: 1087 of Table 2, to the position of nucleic acid number (antisense strand, stop) or a sequence having at least 70 % homology thereto.

15

36. The biomoleculeip to Claim 24, wherein the substrate is addressable.

37. A biomolecule chip with a polypeptide or a variant thereof, having at least one amino acid sequence selected from the group consisting of SEQ ID NO: 2-341, 343-722, 724-1086, 1088-1468, 1470-1837 and 1839-2157, or a sequence having at least 70 % homology thereto, located therein.

38. The biochip according to Claim 37, wherein the polypeptide or the variant thereof, has at least three contiguous amino acid lengths of at least one amino acid sequence selected from the group consisting of SEQ ID NO: 2-341, 343-722, 724-1086, 1088-1468, 1470-1837 and 1839-2157, or a sequence having at least 70 % homology thereto, located therein.

30

39. The biochip according to Claim 37, wherein the

polypeptide or the variant thereof, has at least eight contiguous amino acid lengths of at least one amino acid sequence selected from the group consisting of SEQ ID NO: 2-341, 343-722, 724-1086, 1088-1468, 1470-1837 and  
5 1839-2157, or a sequence having at least 70 % homology thereto, located therein.

40. The biochip according to Claim 37, wherein the polypeptide or the variant thereof, as at least three  
10 contiguous or non-contiguous amino acid lengths of at least one amino acid sequence selected from the group consisting of SEQ ID NO: 2-341, 343-722, 724-1086, 1088-1468, 1470-1837 and 1839-2157, or a sequence having at least 70 % homology thereto, and having a biological function, located therein.

15

41. The biomolecule chip according to Claim 40, wherein the biological activity comprises a function set forth in Table 2.

20

42. The biomolecule chip according to Claim 40, wherein the biological activity comprises epitope activity.

43. A storage medium having stored therein information of a nucleic acid sequence of a nucleic acid molecule having  
25 at least eight contiguous or non-contiguous nucleotides of the sequence set forth in SEQ ID NOs: 1 or 1087, or a variant thereof.

44. The storage medium according to Claim 43, wherein the  
30 nucleic acid molecule or the variant thereof, comprises at least eight contiguous nucleotide lengths of substantially all the sequences encoding sequences selected from the group consisting of SEQ ID NOs: 2-341, 343-722, 724-1086,



1088-1468, 1470-1837 and 1839-2157, or sequences with one or more amino acid substitution, addition and/or deletion thereto.

5    45.    The storage medium according to Claim 43, wherein when  
the reading frame of Table 2 is f-1, f-2 or f-3, the nucleic  
acid molecule or the variant thereof, has a sequence from  
the position of nucleic acid number (sense strand, start)  
of SEQ ID NO: 1 of Table 2, to the position of nucleic acid  
10    number (sense strand, stop) or a sequence having at least  
70 % homology thereto, or when the reading frame of Table  
2 is r-1, r-2 or r-3, the nucleic acid molecule has a sequence  
from the position of nucleic acid number (antisense strand,  
start) of SEQ ID NO: 1087 of Table 2, to the position of  
15    nucleic acid number (antisense strand, stop) or a sequence  
having at least 70 % homology thereto.

46.    A storage medium, comprising information of a  
polypeptide or a variant thereof having at least one amino  
20    acid sequence selected from the group consisting of SEQ ID  
NO: 2-341, 343-722, 724-1086, 1088-1468, 1470-1837 and  
1839-2157, or a sequence having at least 70 % homology  
thereto, located therein.

25    47.    The storage medium according to Claim 46, wherein the  
polypeptide or the variant thereof, has at least three  
contiguous amino acid lengths of at least one amino acid  
sequence selected from the group consisting of SEQ ID NO:  
2-341, 343-722, 724-1086, 1088-1468, 1470-1837 and  
30    1839-2157, or a sequence having at least 70 % homology  
thereto, located therein.

48.    The storage medium according to Claim 46, wherein the

polypeptide or the variant thereof, has at least eight contiguous amino acid lengths of at least one amino acid sequence selected from the group consisting of SEQ ID NO: 2-341, 343-722, 724-1086, 1088-1468, 1470-1837 and  
5 1839-2157, or a sequence having at least 70 % homology thereto, located therein.

49. The storage medium according to Claim 46, wherein the polypeptide or the variant thereof, has at least three  
10 contiguous or non-contiguous amino acid lengths of at least one amino acid sequence selected from the group consisting of SEQ ID NO: 2-341, 343-722, 724-1086, 1088-1468, 1470-1837 and 1839-2157, or a sequence having at least 70 % homology thereto, and having a biological function, located therein.

15

50. The storage medium according to Claim 49, wherein the biological activity comprises a function set forth in Table 2.

20 51. A biomolecule chip having at least one antibody against a polypeptide or a variant thereof, located on a substrate, the polypeptide or the variant thereof comprises at least one amino acid sequence of sequences selected from the group consisting of SEQ ID NOs: 2-341, 343-722, 724-1086,  
25 1088-1468, 1470-1837 and 1839-2157, or a sequence having at least 70 % homology thereto.

52. An RNAi molecule having a sequence homologous to a reading frame sequence wherein, when the reading frame of  
30 Table 2 is f-1, f-2 or f-3, the reading frame sequence has a sequence from the position of nucleic acid number (sense strand, start) of SEQ ID NO: 1 of Table 2, to the position of nucleic acid number (sense strand, stop) or a sequence

having at least 70 % homology thereto, or when the reading frame of Table 2 is r-1, r-2 or r-3, the reading frame sequence has a a sequence from the position of nucleic acid number (antisense strand, start) of SEQ ID NO: 1087 of Table  
5 2, to the position of nucleic acid number (antisense strand, stop) or a sequence having at least 70 % homology thereto.

53. The RNAi molecule according to Claim 52, which is an RNA or a variant thereof comprising a double-stranded  
10 portion of at least 10 nucleotide length.

54. The RNAi molecule according to Claim 52, comprising a 3' overhang terminus.

15 55. The RNAi molecule according to Claim 54, wherein the 3' overhang terminus is a DNA having at least 2 nucleotides in length.

20 56. The RNAi molecule according to Claim 54, wherein the 3' overhang terminus is a DNA of two to four nucleotides in length.